J. Microbiol. Biotechnol. (2007), **17**(8), 1330–1337

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Annotation and Expression Profile Analysis of cDNAs from the Antarctic Diatom *Chaetoceros neogracile*

JUNG, GYEONGSEO[1], CHOUL-GYUN LEE[2], SUNG-HO KANG[3], AND EONSEON JIN[1]*

[1]*Department of Life Science, College of Natural Science, Hanyang University, Seoul 133-791, Korea*
[2]*Department of Biotechnology, Institute of Industrial Biotechnology, Inha University, Incheon 402-751, Korea*
[3]*Polar Environmental Research Division, Korea Ocean Polar Research Institute (KOPRI), Incheon 426-744, Korea*

Received: March 20, 2007 Accepted: May 31, 2007

Abstract To better understand the gene expression of the cold-adapted polar diatom, we conducted a survey of the *Chaetoceros neogracile* transcriptome by cDNA sequencing and expression of interested cDNAs from the Antarctic diatom. A non-normalized cDNA library was constructed from the *C. neogracile*, and a total of 2,500 cDNAs were sequenced to generate 1,881 high-quality expressed sequence tags (ESTs) (accession numbers EL620615-EL622495). Based on their clustering, we identified 154 unique clusters comprising 342 ESTs. The remaining 1,540 ESTs did not cluster. The number of unique genes identified in the data set is thus estimated to be 1,694. Taking advantage of various tools and databases, putative functions were assigned to 939 (55.4%) of these genes. Of the remaining 540 (31.9%) unknown sequences, 215 (12.7%) appeared to be *C. neogracile*-specific since they lacked any significant sequence similarity to any sequence available in the public databases. *C. neogracile* consisted of a relatively high percentage of genes involved in metabolism, genetic information processing, cellular processes, defense or stress resistance, photosynthesis, structure, and signal transduction. From the ESTs, the expression of these putative *C. neogracile* genes was investigated: fucoxanthin chlorophyll (chl) *a,c*-binding protein (*FCP*), ascorbate peroxidase (*ASP*), and heat-shock protein 90 (*HSP90*). The abundance of *ASP* and *HSP90* changed substantially in response to different culture conditions, indicating the possible regulation of these genes in *C. neogracile.*

Keywords: *Chaetoceros neogracile*, Antarctic diatom, ESTs, cold adaptation

More than 70% of the earth consists of cold ecosystems that have a stable temperature below or close to the freezing point of water. Cold habitats include deep ocean, alpine, and polar environments [33]. To successfully colonize low-temperature environments, psychrophilic photoautotrophs have evolved a number of strategies that range from molecular to whole cell to ecosystem levels. The process of genetic change that occurs over a time period of many generations in response to an organism's specific environmental niche is termed adaptation. The Antarctic diatom *Chaetoceros neogracile* is a unicellular photosynthetic eukaryote that is one of the main biomass producers in polar microalgae. One basic prerequisite for thriving in permanent cold environments such as polar seawater and ice is tolerance of, or adaptation to, cold temperatures, which is found in the Antarctic marine diatom *C. neogracile* [33]. In spite of their ecological success in the world's oceans, very little information is available regarding the biology of *C. neogracile* at the molecular level [19].

As a first step to gain insight into the molecular processes occurring in microalgal psychroadaptation, we employed large-scale single-pass sequencing of cDNAs, generally known as expressed sequence tags (ESTs), generated from Antarctic microalgae. ESTs serve as markers for the genes expressed by a certain cell type under specific environmental culture conditions and are used for the recovery of full-length cDNA or genomic clones, discovery of new genes, recognition of exon/intron boundaries, delineation of protein families, and development of genetic maps. Furthermore, ESTs with no homology to known proteins may be used as the first indicators that these unknown proteins are actually expressed. For these reasons, the EST approach is widely used and a large number of ESTs have been collected over the past years [17, 37]. Recently, the first complete genome of an ecologically important marine diatom (*Thalassiosira pseudonana*) was sequenced [5]. The diatom EST database was integrated from two diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* [27]. The database currently contains sequences of close to 30,000 ESTs organized into PtDB (the *P. tricornutum* EST database) and TpDB (the *T.*

*Corresponding author
Phone: 82-2-2220-2561; Fax: 82-2-2299-2561;
E-mail: esjin@hanyang.ac.kr

pseudonana EST database). Approximately 50% of the sequences from newly sequenced genomes and ESTs bear no similarity to genes identified previously [4]. This proportion may be larger for organisms from polar environments, and thus these genes could point to novel physiological and ecological phenomena [32].

The Antarctic diatom *C. neogracile* was selected for molecular studies on the mechanisms of psychrophilic adaptation. The goal of this work was to provide a functionally annotated preliminary set of ESTs from *C. neogracile* expressed under an important polar environmental condition. Our results provide insights into the genome of an Antarctic diatom and provide evidence for specific sequences that may underlie the adaptation to cold temperature conditions in polar seawater. These experiments can increase our general knowledge regarding the physiology of cold adaptation in Antarctic microalgae in general and in *C. neogracile* particularly.

MATERIALS AND METHODS

Culturing Conditions

The unicellular Antarctic marine diatom *C. neogracile* (KOPRI AnM0002) was kindly provided by KOPRI (Korea Ocean Polar Research Institute, Inchon, Korea). The algae were grown in 0.2-μm-filtered natural seawater enriched with F/2 nutrients [12]. The seed culture was grown in a 2-l culture bottle in natural seawater at 4°C under continuous light intensity of 40 μmol photon·m^{-2}·s^{-1}. To investigate the temperature stress response of the Antarctic diatom, the cultures were inoculated at a cell density of 5×10^5 cells/ml with cells in the exponentially growing phase at 4°C. The cultures were moved to a culture chamber at +10°C (±0.5°C) and incubated for the designated time period of the experiment.

cDNA Library Construction

The total RNA was extracted from *C. neogracile* using the TRI reagent (Molecular Research Center, Inc., Cincinnati, OH, U.S.A.). Total RNA of each sample was quantified spectrophotometrically at OD_{260}. An OD_{260} of 1 corresponded to 40 μg/ml RNA. Subsequently, the RNA was precipitated and resuspended in DEPC-treated distilled water to a final concentration of 1 mg/ml. Poly(A)+ RNA was then prepared using a Poly(A) Track mRNA isolation system (Promega, Madison, WI, U.S.A.) and a cDNA library constructed using the ZAP-cDNA synthesis kit and the ZAP-cDNA Gigapack III Gold packaging extract (Stratagene, La Jolla, CA, U.S.A.) according to the manufacturer's instructions. The total primary titer of each library in recombinant plaque-forming units was 6.5×10^6. After amplification, samples from each cDNA library were used to subclone inserts by mass excision for the conversion of Lambda to a phagemid vector. The resulting phagemid libraries were plated at low density on Luria Bertani agar plates containing kanamycin (25 mg/l).

Nucleotide Sequencing

A total of 2,500 colonies were randomly picked and cultivated for storage and isolation of the plasmid DNA. The plasmid DNA was purified from the *E. coli* cultures by alkaline lysis, vacuum filtration and anion-exchange chromatography using a high-throughput, 96-well format system (Millipore, MA, USA). The amplified cDNAs were purified with GeneClean (BIO 101,Vista, CA, U.S.A.) and the 5'-end from each cDNA was sequenced with the T3 primer and ABI Prism BigDye Terminator Cycle Sequencing reaction mix (PE Applied Biosystems, Foster City, CA, U.S.A.). The reactions were purified using Centri-Sep spin columns (Princeton Separations, Adelphia, NJ, U.S.A.) and the eluted extension products were resolved and analyzed on an ABI 3100 Genetic Analyzer (PE Applied Biosystems).

Analysis of cDNAs and Functional Annotation

Prior to annotation, the raw DNA sequence data were edited to remove the vector sequences and poor quality data using a computer program (Lasergene software, DNASTAR, Inc., Madison, WI, U.S.A.). The edited EST sequences were then arranged into nucleotide-matched clusters using an alignment program (SeqMan in Lasergene, DNASTAR, Inc.) to determine the frequency of sampling redundancy. The criterion used for a sequence to be identified as nonredundant was based on a unique best match based on the accession numbers from the BlastX search. The accession numbers and *E* values of the best matches were extracted from the result file. The *E* value is a parameter that describes the number of hits one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially with the score (S) that is assigned to a match between two sequences. If two or more query sequences resulted in best matches with identical accession numbers, they were sorted according to their *E* values. Assignment of putative identities for ESTs required a minimum *E* value of 10^{-4}. ESTs with known gene matches were categorized into different functional groups according to categories described in Eom *et al*. [17]. Each EST that did not meet this requirement was annotated as unknown. The complete list of sequences was deposited in the NCBI EST database (http://www.ncbi.nlm.nih.gov/projects/dbEST) and is available with the following accession numbers: EL620615-EL622495

Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

RT-PCR was performed on total RNA prepared from *C. neogracile* incubated for 0, 5, 15, and 30 min, at +10°C

Table 1. Primer sequences for RT-PCR and their expected product size.

Genes	Forward (F) and reverse (R) primer sequences	Product size (bp)
Fucoxanthin chlorophyll a,c-binding protein (FCP)	F : 5'-CCCAGCCTTCAGCAGCGTC-3' R : 5'-CCCAGCCTTCAGCAGCGTC-3'	507
Ascorbate peroxidase (ASP)	F : 5'-GCATCATCAACTTCCCAGTAG-3' R : 5'-GGACGATACCACATCCCAAGC-3'	530
Heat shock protein 90 (HSP 90)	F : 5'-CGGGCAAAGTTGAGGCAGGC-3' R : 5'-GCAAGACGCTCAAGAAGAAGAG-3'	494
Tubulin (TUB)	F : 5'-GAGCATTCACATCGGTCAGGC-3' R : 5'-GCAAGACGCTCAAGAAGAAGAG-3'	459

and 4°C (control) using the One-Step RT-PCR system (Invitrogene, CA, U.S.A.). To amplify the different *C. neogracile* genes, the gene specific primers were designed. Table 1 summarizes the sources and sequences of primers used in this study. cDNA synthesis and pre-denaturation were performed in single cycles at 45°C for 40 min and 94°C for 2 min and PCR amplification was performed for 35 cycles at 94°C for 30 sec, 55°C for 45 sec, and 72°C for 1 min. The final elongation step was performed at 72°C for 10 min in MyCycler™ Thermal Cycler (BIO-RAD, CA, U.S.A.). The PCR products obtained were separated on 1% agarose gels, stained with ethidium bromide and documented in a gel documentation system (Vilber Lourmat, France). To normalize the RT-PCR data, each gene was compared with tubulin transcript, whose expression was constant under all culture conditions.

Results and Discussion

Overview of ESTs from *C. neogracile* cDNA Library

The cDNA library from the Antarctic diatom *C. neogracile* was created in a Uni-Zap XR vector (Stratagene) using oligo-dT primers and directionally inserted into EcoRI-XhoI sites of pBluescript. The 5'-end sequences were generated using the T3 primer. A total of 2,500 random clones were partially sequenced from this cDNA library to generate ESTs. The edited 1,881 EST sequences were then arranged into nucleotide-matched clusters using an alignment program (SeqMan in Lasergene, DNASTAR) to determine the frequency of sampling redundancy. These ESTs were clustered and assembled into a collection of 1,694 nonredundant sequences with 154 contigs and 1,540 singleton ESTs (Table 2).

Table 2. Summary of ESTs from Antarctic diatom *C. neogracile*.

Number of clones	Number of contigs	Number of clones
Single	1,540	1,540
Group	154	342
Total	1,694	1,881

Annotation and Functional Classification

Each gene in the nonredundant EST groups inherited the annotation from the best match found after a BlastX search against the nr protein database at NCBI. An expectation value (*E* value) threshold of 10^{-4} was used. Comparing the similarity results against a public database, a total of 939 (55.4%) of the 1,694 nonredundant EST groups were found to be similar to genes registered in the databases. Other 540 EST sequences (with *E* values $>10^{-4}$ in the Blastx search versus nr) were annotated as unknown. The remaining 215 ESTs (12.7%) showed no similarity to the sequences of registered genes and were classified as novel sequences (Table 3). Overall, almost half (44.6%) of the tentatively identified genes expressed in the Antarctic diatom had no known function and no similarity to known genes, and were therefore a potentially significant resource for identifying new genes. In Fig. 1, the functional category is indicated by the text associated with the respective piece in the diagram. The size of each piece is proportional to the relative abundance of the proteins assigned to this group. *C. neogracile* consisted of a relatively high percentage of genes involved in metabolism (16.6%), followed by genes involved in genetic information processing (7.6%), photosynthesis (4.2%), defense and stress resistance (2.4%), cell signaling and communication (2.3%), transport (2.2%), and cell structure (0.7%). Unclassified genes constituted the remaining 16.2%. These functional expression profiles are summarized in Table 4.

Highly Expressed Genes from the Antarctic Marine Diatom *C. neogracile*

To gain insight into the major biological processes in the Antarctic diatom *C. neogracile*, the activities of the 25

Table 3. Results of a similarity search with the public database. The number of EST groups and clones exhibiting a similarity to genes with known functions and unknown genes with no similarity match with public data are given.

	Group	(%)
Gene with known function	939	55.4
Unknown	540	31.9
No similarity	215	12.7
Total	1,694	100

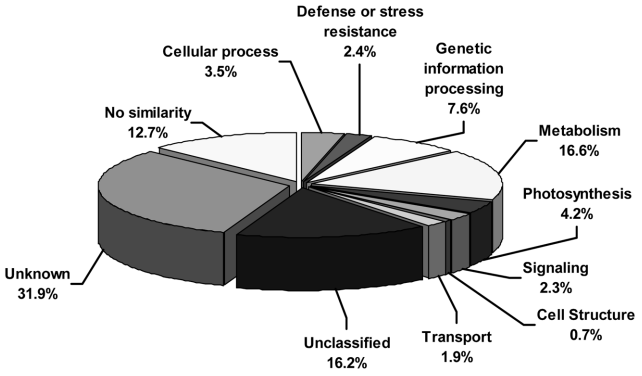

Fig. 1. Expression analysis of the 1,881 ESTs from *C. neogracile*. No similarity and no sequence similarity were found among known amino acid or DNA sequences.

Table 4. Classification of the ESTs according to putative function.

Catalog	Number	(%)
Cellular process	60	3.5
Defense or stress resistance	41	2.4
Genetic information processing	128	7.6
Metabolism	281	16.6
Photosynthesis	71	4.2
Signaling	39	2.3
Cell Structure	12	0.7
Transport	33	1.9
Unclassified	274	16.2
Unknown	540	31.9
No similarity	215	12.7
Total	1,694	100

most highly expressed genes were examined, which accounted for 18% of nonredundant *C. neogracile* ESTs. To overcome cold temperatures and survive in extreme environments, this strain required energy and metabolism from carbohydrates and adaptation to extreme stress. These genes encoded one or more protein turnover functions and putative defense proteins (Table 5). The protein mostly detected was FCP. A total of 80 ESTs

Table 5. Top 25 genes represented in the *C. neogracile*.

Rank	Best BlastX Hit Description	*E*-value*	No. of ESTs
1	Fucoxanthin chlorophyll *a/c* protein	7E-69	80
2	Plasma membrane H+-ATPase	1E-04	16
3	Ubiquitin extension protein	1E-61	15
4	Sinapryl alcohol dehydrogenase	2E-05	14
5	ThiJ/PfpI	3E-53	12
6	Eukaryotic initiation factor 4A	1E-131	11
6	Glyceraldehyde-3-phosphate dehydrogenase	2E-114	11
8	40S ribosomal protein	4E-85	10
8	Glutathione S-transferase, C-terminal-like	3E-58	10
8	Oxidoreductase/zinc ion binding	1E-73	10
11	60S acidic ribosomal protein PO	1E-40	9
11	Cell division protease ftsH homolog, chloroplast precursor	3E-82	9
13	4-Coumaryl-CoA ligase	2E-04	8
13	Alcohol dehydrogenase, zinc-binding domain protein	2E-28	8
15	Elongation factor 1 alpha long form	1E-86	7
15	Transaldolase	1E-88	7
17	Actin	5E-124	6
17	Ankyrin repeat family protein /methyltransferase-related	3E-26	6
17	ATP synthase alpha-subunit	2E-134	6
17	Protein kinase	2E-38	6
21	6-Phosphogluconate dehydrogenase	1E-53	5
21	Acyl-CoA dehydrogenase	3E-65	5
21	Heat-shock protein 90	2E-70	5
21	Ubiquitin activating enzyme E1	1E-28	5
25	ABC transporter	4E-48	4
25	ADP,ATP carrier protein	4E-73	4
25	Beta-galactosidase	5E-30	4
25	Fructose-1,6-bisphosphatase	3E-69	4
25	Diaminopimelase decarboxylase	2E-29	4
25	Glutathione reductase	4E-26	4

**E*-value is Expect value.

encode FCP in *C. neogracile*. The major light-harvesting complex in diatoms is FCP, which can transfer energy to chl *a* of the photosynthetic reaction centers [20]. FCPs range in mass from 18 to 19.5 kDa [9]. Previously, some FCPs were reported to be highly expressed in *F. cylindrus* and *P. tricornutum* [30, 31, 44]. Recently, a number of FCP genes, with high sequence homology, have been characterized in several species of fucoxanthin-containing algae [9, 12, 16, 18, 23]. Our results showing that a large number of this gene family (*FPCs*) being highly expressed under the present culture condition indicated a large amount of energy transfer to photosystems may be required to provide enough energy for photosynthesis in a polar environment.

6-Phosphogluconate dehydrogenase (6PGDH) was also found to be one of the abundant ESTs from *C. neogracile* (Table 5). 6PGDH catalyzes the $NADP^+$-dependent oxidative decarboxylation of 6-phosphogluconate to ribulose-5-phosphate. With glucose-6-phosphate dehydrogenase (6-GDH), 6PGDH in the Antarctic diatom appears to be an important protein providing NADPH for reductive biosyntheses and for protection against oxidative stress, pentoses for synthesis of nucleotides and sugar phosphates for the shikimate pathway, and involving in the adaptation of several chloroplast-cytosol isoenzymes to the cold environment.

Several genes that encode proteins implicated in playing a role in stress response or defense are highly expressed in *C. neogracile*. Among these, the most highly represented is the family of GSTs, (EC 2.5.1.18). Previously, GSTs from a polar diatom had never been reported. In contrast, plant GSTs are known to be encoded by a large and diverse gene family [15]. The inducibility of some GSTs following exposure of plants to biotic and abiotic stresses is a characteristic feature of these genes [28]. In some instances it appears likely that GSTs are induced by the general oxidative stress caused by these diverse treatments, but in other cases, GST induction is specific to the particular stress [29]. Expression of GSTs is also enhanced by a range of xenobiotics and osmotic stress and extreme temperatures [28]. The existence of additional functions of GSTs, not directly derived from their ability to catalyze the formation of GSH conjugates, has gained further support with studies demonstrating that several stress-inducible GSTs protect plants from oxidative injury by functioning as glutathione peroxidases [14, 41]. Certain GSTs have been shown to have glutathione peroxidase activity, with the GSTs using glutathione to reduce organic hydroperoxides of fatty acids and nucleic acids to the corresponding monohydroxyalcohols. This reduction plays a pivotal role in preventing the degradation of organic hydroperoxides to cytotoxic aldehyde derivatives. This functionality in GSTs has been demonstrated to be important in the tolerance of transgenic tobacco to chilling and salt [22, 41]. Glutathione, the substrate of GST, has numerous roles in cellular defense and in sulfur metabolism. Glutathione reacts chemically with a range of active oxygen species (AOS), whereas enzyme-catalyzed reactions link GSH to the detoxification of H_2O_2 in the ascorbate-glutathione cycle. Importantly, GSH protects proteins against the denaturation that is caused by oxidation of protein thiol groups during stress. All these functions involve the oxidation of the thiol group, principally to form glutathione disulfide (GSSG). Cellular GSH:GSSG ratios are maintained by glutathione reductase (GR, EC 1.6.4.2), a homodimeric flavoprotein that uses NADPH to reduce GSSG to two GSHs [35]. The redox state of glutathione will depend on the balance between oxidative processes and the *in vivo* GR activity [35]. The Antarctic diatom in this study has four putative GR ESTs (Table 5). Polar diatoms may require extra GSTs and GRs to handle extreme environments such as low temperature, UV light, and high salinity to which they are easily exposed.

Chaperone or Protein Turnover Function

FtsH (filamentation temperature-sensitive) proteases are ATP- and zinc-dependent metallo-type proteases, which are highly expressed in *C. neogracile*. The proteases mediate degradation of membrane proteins in mitochondria and chloroplasts [25]. In chloroplasts, the FtsH proteases are anchored to thylakoid membranes with two transmembrane segments at the N terminus and a large C terminus part exposed to the stroma [26]. FtsH forms a hexameric ring, and substrate proteins are translocated through a central cavity in an ATP-dependent manner. The substrates are subsequently digested at the zinc-binding site. In the chloroplasts of *Arabidopsis*, 12 FtsH proteases (FtsH1-FtsH12) are present; in particular, FtsH2 and FtsH5 are shown to participate in the repair of photodamaged PS II by digestion and removal of the damaged D1 protein under light stress conditions [7, 34, 42, 43]. The turnover process of the photodamaged PS II under light stress has been studied extensively [6, 8, 24, 36, 47], and according to these results, monomerization of the PS II dimer and regulation of turnover of the D1 protein by protein phosphorylation/dephosphorylation are crucial for recognition and degradation of the damaged D1 protein by a specific protease(s) [48]. This relatively abundant transcript found in the Antarctic diatom may need further investigation for the function of this gene.

A proportionally high number of HSP90 transcripts were detected in the *C. neogracile*. The HSP90 family represents a highly conserved chaperone class. Heat-shock proteins can also be induced by environmental stress. HSP90 members are found in the cytosol, the endoplasmic reticulum, mitochondria, and chloroplasts of eukaryotic cells [13, 46]. The functional role of the HSP90 in *C. neogracile* requires further investigation to discern the cold adaptation mechanism.

Expression of the *C. neogracile* Genes

To explore whether the identified diatom genes are regulated under temperature changes, an expression analysis of the genes of interest was performed. RT-PCR was run on total RNA isolated from Antarctic diatom, with temperature change (to +10°C) at time points between 5 min to 30 min using gene-specific primers for putative *C. neogracile ASP*, *HSP90*, and *FCP*, respectively. Total RNA isolated from untreated diatom at the same time points was used for comparison. A diatom *TUB* was also amplified from the same RNA samples as a loading and RNA quality control. The putative *ASP* gene was weakly expressed at time 0, but was rapidly induced at the first time point after 5 min. The expression levels continued thereafter to increase and peaked after approximately 15 min but still showed an elevated expression at 30 h. The putative *HSP90* gene was slowly induced up until 30 min. A majority of the known plant HSP90s are abundant cellular proteins whose concentration is increased up to 10-fold in response to stress [10]. The major function attributed to cytosolic HSP90 is a role in the maturation of signal transduction proteins, such as hormone receptors and kinases [39, 40, 45]. Therefore, the upregulation of the *HSP90* under temperature change in the Antarctic diatom is consistent with the results of previous studies [1, 21], in which cytosolic HSP90 participated in the regulation of the stress response. However, the *FCP* gene expression remained at the same level under 10°C cultivation (Fig. 2). The *FCP* was particularly interesting since this gene was expressed constitutively under heat stress. Expression of these *FCP* genes is known to be influenced by the light intensity [3, 38] and the rate of transcription of individual genes is differentially modulated by blue light and light intensity [2]. *FCP* may have been resistant up to 30 min by high temperature stress conditions where other genes (*ASP* and *HSP90*) were sensitive. The results indicate that the putative *ASP* and *HSP90* genes have intricate and different individual roles in response to temperature change and that they adjust to extreme environmental conditions for growth.

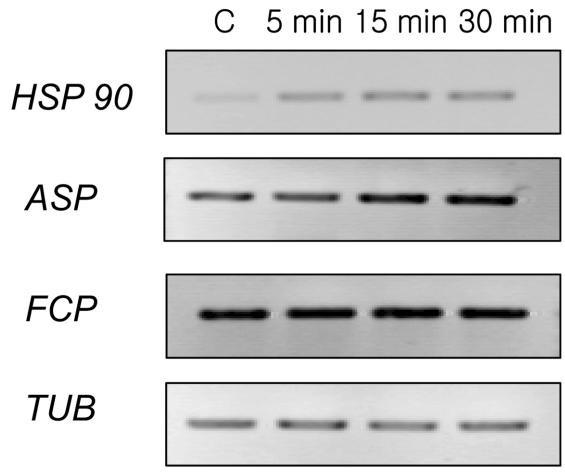

Fig. 2. RT-PCR analysis of putative *FCP*, *ASP*, and *HSP90* expression in *C. neogracile*.
Diatoms were incubated at +10°C for the times indicated below the lanes (c: control; min: minutes). Total RNA was isolated from leaves and PCR reactions were run in 35 cycles using putative *C. neogracile FCP*, *ASP*, or *HSP90* specific primers, as indicated to the left of the picture. Equal loading and RNA quality were controlled by amplification of the *TUB* gene.

In conclusion, the characterization of 1,881 ESTs from a cold-adapted cDNA library was reported. This collection of 1,881 ESTs provides a preliminary view into the gene expression profile of the Antarctic diatom *C. neogracile*. We also investigated how the expression of the identified genes in the Antarctic diatom related to temperature changes. This EST database represents a new genomic tool for studying aspects of cold-adapted mechanisms at the level of gene expression and is a resource for novel candidate gene discovery. The *C. neogracile* EST data set can now be used to fabricate a cDNA chip, to perform various expression studies at varying temperatures, and should provide valuable insights into the molecular mechanisms of cold adaptation.

Acknowledgment

This work was supported by grant No. R01-2006-000-10856-0 from Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST), in part by KOPRI PE07060 to S.H. Kang.

REFERENCES

1. Ali, A., S. Bharadwaj, R. O'Carroll, and N. Ovsenek. 1998. HSP90 interacts with and regulates the activity of heat shock factor 1 in *Xenopus* ocytes. *Mol. Cell. Biol.* **18:** 4949–4960.
2. Apt, K. E., S. K. Clendennen, D. A. Powers, and A. R. Grossman. 1995. The gene family encoding the fucoxanthin chlorophyll proteins from the brown alga *Macrocystis pyrifera*. *Mol. Gen. Genet.* **246:** 455–464.
3. Apt, K., D. Bhaya, and A. R. Grossman. 1994. Characterization of genes encoding the light-harvesting proteins in diatoms: Biogenesis of the fucoxanthin chlorophyll *a/c* protein complex. *J. Appl. Phycol.* **6:** 225–230.
4. Armbrust, E. V., J. A. Berges, C. Bowler, B. R. Green, D. Martinez, N. H. Putnam, S. Zhou, A. E. Allen, K. E. Apt, M. Bechner, M. A. Brzezinski, B. K. Chaal, A. Chiovitti, A. K. Davis, M. S. Demarest, J. C. Detter, T. Glavina, D. Goodstein, M. Z. Hadi, U. Hellsten, M. Hildebrand, B. D. Jenkins, J. Jurka, V. V. Kapitonov, N. Kröger, W. W. Lau, T. W. Lane, F. W. Larimer, J. C. Lippmeier, S. Lucas, M. Medina, A. Montsant, M. Obornik, M. S. Parker, B. Palenik, G. J. Pazour, P. M. Richardson, T. A. Rynearson, M. A. Saito,

D. C. Schwartz, K. Thamatrakoln, K. Valentin, A. Vardi, F. P. Wilkerson, and D. S. Rokhsar. 2003. The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science* **306:** 79–86.

5. Armbrust, E. V., J. B. Berges, C. Bowler, B. R. Green, D. Martinez, N. H. Putnam, S. Zhou, A. E. Allen, K. E. Apt, M. Bechner, *et al.* 2004. The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science* **306:** 79–86.
6. Aro, E.-M., I. Virgin, and B. Andersson. 1993. Photoinhibition of photosystem II. Inactivation, protein damage and turnover. *Biochim. Biophys. Acta* **1143:** 113–134.
7. Bailey, S., E. Thompson, P. J. Nixon, P. Horton, C. W. Mullineaux, C. Robinson, and N. H. Mann. 2002. A critical role for the Var2 FtsH homologue of *Arabidopsis thaliana* in the photosystem II repair cycle *in vivo*. *J. Biol. Chem*. **277:** 2006–2011.
8. Barber, J. and B. Andersson. 1992. Too much of a good thing: Light can be bad for photosynthesis. *Trends Biochem. Sci.* **17:** 61–66.
9. Bhaya, D. and A. R. Grossman. 1993. Characterization of gene clusters encoding the fucoxanthin chlorophyll proteins of the diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res.* **21:** 4458–4466.
10. Buchner, J. 1999. Hsp90 & Co. - a holding for folding. *Trends Biochem. Sci*. **24:** 136–141.
11. Caron, L., D. Douady, M. Quinet-Szely, S. de Goër, and C. Berkaloff. 1996. Gene structure of a chlorophyll *a/c*-binding protein from a brown alga: Presence of an intron and phylogenetic implications. *J. Mol. Evol.* **43:** 270–280.
12. Ceron, G., F. Garcia CamaCho, A. Sanchez Miron, J. M. Fernandez Sevilla, Y. Chisti, and E. M. Grima. 2006. Mixotrophic production of marine microalga *Phaeodactylum tricornutum* on various carbon sources. *J. Microbiol. Biotechnol*. **16:** 689–694.
13. Csermelyl, P. and S. C. Lakhotia. 1998. Molecular biology of stress responses in India. *Cell Stress Chaperones* **3:** 1–5.
14. Cummins, I., D. J. Cole, and R. Edwards. 1999. A role for glutathione transferases functioning as glutathione peroxidases in resistance to multiple herbicides in black-grass. *Plant J.* **18:** 285–292
15. Dixon, D. P., A. Lapthorn, and R. Edwards. 2002. Plant glutathione transferases. Protein family review. *Genome Biol*. **3:** reviews 3004.1–3004.10.
16. Durnford, D. G, R. Aebersold, and B. R. Green. 1996. The fucoxanthin-chlorophyll proteins from a chromophyte alga are part of a large multigene family: Structural and evolutionary relationships to other light harvesting antennae. *Mol. Gen. Genet.* **253:** 377–386.
17. Eom, H. S., S. H. Park, C. G. Lee, and E. S. Jin. 2005. Gene expression profiling of eukaryotic microalga, *Haematococcus pluvialis*. *J. Microbiol. Biotechnol*. **15:** 1060–1066.
18. Eppard, M. and E. Rhiel. 1998. The genes encoding light-harvesting subunits of *Cyclotella cryptica* (*Bacillariophyceae*) constitute a complex and heterogeneous family. *Mol. Gen. Genet*. **260:** 335–345.
19. Falciatore, A. and C. Bowler. 2002. Revealing the molecular secrets of marine diatoms. *Annu. Rev. Plant Biol*. **53:** 109–130.
20. Grossman, A. R., A. Manodori, and D. Snyder. 1990. Light-harvesting proteins of diatoms: Their relationship to the chlorophyll *a/b* binding proteins of higher plants and their mode of transport into plastids. *Mol. Gen. Genet.* **224:** 91–100.
21. Jakob, U., I. Meyer, H. Bügl, S. André, J. C. Bardwell, and J. Buchner. 1995. Structural organization of prokaryotic and eukaryotic Hsp90. Influence of divalent cations on structure and function. *J. Biol. Chem*. **270:** 14412–14419.
22. Kampranis, S. C., R. Damianova, M. Atallah, G. Toby, G. Kondi, P. N. Tsichlis, and A. M. Makris. 2000. A novel plant glutathione S-transferase/peroxidase suppresses Bax lethality in yeast. *J. Biol. Chem*. **275:** 29207–29216.
23. Kroth-Pancic, P. 1995. Nucleotide sequence of two cDNAs encoding fucoxanthin chlorophyll *a/c* proteins in the diatom *Odontella sinensis*. *Plant Mol. Biol*. **27:** 825–828.
24. Kyle, D. J., I. Ohad, and C. J. Arntzen. 1984. Membrane protein damage and repair: Selective loss of a quinone-protein function in chloroplast membranes. *Proc. Natl. Acad. Sci. USA* **81:** 4070–4074.
25. Langer, T. 2000. AAA proteases: Cellular machines for degrading membrane proteins. *Trends Biochem. Sci.* **25:** 247–251.
26. Lindahl, M., S. Tabak, L. Cseke, E. Pichersky, B. Andersson, and Z. Adam. 1996. Identification, characterization, and molecular cloning of a homologue of the bacterial FtsH protease in chloroplasts of higher plants. *J. Biol. Chem*. **271:** 29329–29334.
27. Maheswari, U., A. Montsant, J. Goll, S. Krishnasamy, K. R. Rajyashri, V. M. Patell, and C. Bowler. 2005. The Diatom EST Database. *Nucleic Acids Res*. **33:** 344–347.
28. Marrs, K. A. 1996. The functions and regulation of glutathione S-transferases in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol*. **47:** 127–158.
29. Mauch, F. and R. Dudler. 1993. Differential induction of distinct glutathione S-transferases of wheat by xenobiotics and by pathogen attack. *Plant Physiol*. **102:** 1193–1201.
30. Mock, T. and K. Valentin. 2004. Photosynthesis and cold acclimation: Molecular evidence from a polar diatom. *J. Phycol*. **40:** 732–741.
31. Mock, T. and N. Hoch. 2005. Long-term temperature acclimation of photosynthesis in steady-state cultures of the polar diatom *Fragilariopsis cylindrus*. *Photosynth. Res*. **85:** 307–317.
32. Mock, T., A. Krell, G. Glöckner, Ü. Kolukisaoglu, and K. Valentin. 2005. Analysis of expressed sequence Tag (ESTs) from the polar diatom *Fracilariopsis cylindrus*. *J. Phycol*. **42:** 78–85.
33. Morgan-Kiss, R. M., J. C. Priscu, T. Pocock, L. Gudynaite-Savitch, and N. P. Huner. 2006. Adaptation and acclimation of photosynthetic microorganisms to permanently cold environments. *Microbiol. Mol. Biol. Rev.* **70:** 222–252.
34. Nixon, P. J., M. Baker, M. Boehm, R. de Vries, and J. Komenda. 2005. FtsH-mediated repair of the photosystem II

complex in response to light stress. *J. Exp. Bot.* **56:** 357–363.
35. Noctor, G., L. Gomez, H. Vanacker, and C. H. Foyer. 2002. Interactions between biosynthesis, compartmentation and transport in the control of glutathione homeostasis and signaling. *J. Exp. Bot.* **53:** 1283–1304.
36. Ohad, I., D. J. Kyle, and C. J. Arntzen. 1984. Membrane protein damage and repair: Removal and replacement of inactivated 32-kilodalton polypeptides in chloroplast membranes. *J. Cell Biol.* **270:** 14919–14927.
37. Park, S. H., J. E. W. Polle, A. Melis, T. K. Lee, and E. S. Jin. 2006. Up-regulation of photoprotection and PSII-repair gene expression by irradiance in the unicellular green alga *Dunaliella salina*. *Marine Biotechnol.* **8:** 120–128.
38. Passaquet, C. and C. Lichtlé. 1995. Molecular study of a light-harvesting apoprotein of *Giraudyopsis stellifer* (Chrysophyceae). *Plant Mol. Biol.* **29:** 135–148.
39. Pratt, W. B. and D. O. Toft. 2003. Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery. *Exp. Biol. Med. (Maywood).* **228:** 111–133.
40. Richter, K. and J. Buchner. 2001. Hsp90: Chaperoning signal transduction. *J. Cell. Physiol.* **188:** 281–290.
41. Roxas, V. P., R. K. Smith, E. R. Allen, and R. D. Allen. 1997. Overexpression of glutathione S-transferase/glutathione peroxidase enhances the growth of transgenic tobacco seedlings during stress. *Nat. Biotechnol.* **15:** 988–991.
42. Sakamoto, W., A. Zaltsman, Z. Adam, and Y. Takahashi. 2003. Coordinated regulation and complex formation of yellow variegated1 and yellow variegated2, chloroplastic FtsH metalloproteases involved in the repair cycle of photosystem II in *Arabidopsis* thylakoid membranes. *Plant Cell* **15:** 2843–2855.
43. Sakamoto, W., T. Tamura, Y. Hanba-Tomita, and M. Murat. 2002. The VAR1 locus of *Arabidopsis* encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. *Genes Cells* **7:** 769–780.
44. Scala, S., N. Carels, A. Falciatore, M. L. Chiusano, and C. Bowler. 2002. Genome properties of the diatom *Phaeodactylum tricornutum*. *Plant Physiol.* **129:** 993–1002.
45. Wegele, H., L. Muller, and J. Buchner. 2004. Hsp70 and Hsp90 - a relay team for protein folding. *Rev. Physiol. Biochem. Pharmacol.* **151:** 1–44
46. Willmund, F. and M. Schroda. 2005. Heat shock protein 90C is a bona fide Hsp90 that interacts with plastidic HSP70B in *Chlamydomonas reinhardtii*. *Plant Physiol.* **138:** 2310–2322.
47. Yamamoto, Y. 2001. Quality control of photosystem II. *Plant Cell Physiol.* **42:** 121–128.
48. Yoshioka, M., S. Uchida, H. Mori, K. Komayama, S. Ohira, N. Morita, T. Nakanishi, and Y. Yamamoto. 2006. Quality control of photosystem II. Cleavage of reaction center D1 protein in spinach thylakoids by FtsH protease under moderate heat stress. *J. Biol. Chem.* **281:** 21660–21669.